DOCUMENT RESUME

ED 166 210 TM 007 805

AUTHOR Powers, James E.
TITLE [A Bayesian Solution for 2 to the k power Factorial Arrangements of Treatments].
PUB DATE 77
NOTE 15p.; Paper presented at the Annual Meeting of the Northeastern Educational Research Association (Ellenville, New York, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Analysis of Variance; *Bayesian Statistics; Hypothesis Testing; Tests of Significance

ABSTRACT

A Bayesian analysis for 2 to the k power factorial arrangements of treatments is presented in this paper. To perform the analysis, an experimenter must specify prior distributions on an orthogonal set of linear functions representing the main effects and interactions and on a function representing the grand mean. The solution is relatively straightforward and not unduly complex numerically. The independence of the functions and the fact that they represent the primary concerns in a factorial arrangement aids in the conceptualization of prior distributions. The independence also facilitates the computation of likelihoods since univariate normal distributions must be evaluated rather than a multivariate normal distribution. An experimenter may desire posterior distributions on other comparisons of the treatment means that are not included in the orthogonal set. For these situations a method is presented whereby posterior distributions for any Scheffe type contrast of the means can be generated without great difficulty. Although the solution given in this paper is specifically for 2 to the k power factorial arrangements, the method is immediately generalizable to any factorial. (Author/CTM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED166210

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Bayesian Solution for 2^k Factorial Arrangements of Treatments

James E. Powers

State University of New York at Albany

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

James E. Powers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

TM007 805

Presented at the 1977 meeting of the Northeastern Educational Research Association, Ellenville, New York.

A Bayesian Solution for 2^k Factorial Arrangements of Treatments

James E. Powers

State University of New York at Albany

Abstract

A Bayesian analysis for 2^k factorial arrangements of treatments is presented in this paper. To perform the analysis an experimenter must specify prior distributions on an orthogonal set of linear functions representing the main effects and interactions and on a function representing the grand mean. The solution is relatively straightforward and not unduly complex numerically. The independence of the functions and the fact that they represent the primary concerns in a factorial arrangement aids in the conceptualization of prior distributions. The independence also facilitates the computation of likelihoods since univariate normal distributions must be evaluated rather than a multivariate normal distribution. An experimenter may desire posterior distributions on other comparisons of the treatment means that are not included in the orthogonal set. For these situations a method is presented whereby posterior distributions for any Scheffe type contrast of the means can be generated without great difficulty. Although the solution given in this paper is specifically for 2^k factorial arrangements, the method is immediately generalizable to any factorial.

A Bayesian Solution for 2^k Factorial Arrangements of Treatments

James E. Powers

State University of New York at Albany

INTRODUCTION

Bayesian statistical methods have not yet been extensively applied to the analysis of variance. Two reasonably complete solutions have been given for the completely random, or one-way, classification. Lindley (1965) gives an analysis of this design including the usual overall test of treatment means and the equivalent of linear contrasts among the means when the prior distributions for all means and the natural logarithm of the unknown population variance are uniform and independent. Powers (1971, 1973) has shown a solution for this situation, regardless of the prior distributions selected, which employs planned orthogonal and Scheffé comparisons of the treatment means.

Several writers have applied Bayesian methods to multiple comparison procedures. For example, Duncan (1961, 1965) and Waller and Duncan (1969), have studied Bayes decision rules. They show that a Bayes rule can be obtained for any prior distribution when pairs of means are analyzed and propose the use of a type 1 to type 2 error weight ratio instead of the usual significance level for multiple comparisons. Box and Tiao (1968) have shown a Bayesian method to estimate treatment means

in the one way random effects model using locally uniform and independent priors.

To date, the most complex design analyzed within the Bayesian framework is the two-way classification. Lindley (1972) presented this solution which is an extension of a general theory presented by Lindley and Smith (1972) whereby estimates of cell means are obtained that depend on the relative sizes of the main effects and the interaction. The solution requires that a diffuse prior be placed on the overall mean and that the likelihood distribution be multivariate normal. The posterior distributions thus generated are then also multivariate normal. The usual multiple comparison procedures are rendered unnecessary by this approach.

The purpose of the present paper is to present a Bayesian solution for the 2^k series of factorial arrangements of treatments. The solution is an extension of that given by Powers (1971, 1973) for the completely random design.

RATIONALE

The basis for this solution is the fact that a factorial arrangement of treatments is simply a special case of the completely random design where the treatments consist of all possible combinations of the levels of each factor in the design.

When a 2^k factorial is employed, the primary interests of the researcher are usually the main effects and the interactions. These can be expressed as a set of k-1 orthogonal linear

contrasts of the treatment means. Powers (1973) shows that if a set of orthogonal linear contrasts are specified on the treatment means, then, due to the independence obtained, a prior distribution may be placed on each contrast, the likelihoods evaluated and the posterior derived for that contrast, independent of each other contrast. Thus, the required Bayesian analysis is greatly simplified since independent priors (rather than a joint prior) may be placed on each of the contrasts and univariate normal distributions (rather than a multivariate normal distribution) may be evaluated to obtain likelihoods. This, of course, makes the form of the prior immaterial since likelihoods can be obtained from a table of the standard normal distribution and the required posterior is rather easily obtained numerically.

REQUIRED ORTHOGONAL CONTRASTS

For the sake of simplicity the contrasts necessary for a 2^3 factorial will be derived here. Generalization of the method to any other 2^k factorial is immediate.

It is assumed that there are an equal number of observations in each treatment combination and that they are normally distributed. In addition, it is assumed that homoscedasticity of variance prevails across treatments and that this population variance (σ^2) is known. In most instances where σ^2 is unknown the error mean square may be used to replace it in the analysis (Hughes and Powers, 1971).

Let A, B and C represent the three factors in the design and

a_1, a_2; b_1, b_2; and c_1, c_2 represent the lower and upper levels, respectively, of these factors. The treatment combinations thus defined are

(1)	$a_1b_1c_1$	(5)	$a_2b_1c_1$
(2)	$a_1b_1c_2$	(6)	$a_2b_1c_2$
(3)	$a_1b_2c_1$	(7)	$a_2b_2c_1$
(4)	$a_1b_2c_2$	(8)	$a_2b_2c_2$

Denote the sample means of these combinations as $\bar{X}_i$ (i = 1 to 8) and the corresponding population means as μ_i (i = 1 to 8).

A set of orthogonal linear contrasts of the means of these treatments representing the main effects and interactions in the population is now

$$\lambda_A = \mu_5 + \mu_6 + \mu_7 + \mu_8 - \mu_1 - \mu_2 - \mu_3 - \mu_4$$

$$\lambda_B = \mu_3 + \mu_4 + \mu_7 + \mu_8 - \mu_1 - \mu_2 - \mu_5 - \mu_6$$

$$\lambda_C = \mu_2 + \mu_4 + \mu_6 + \mu_8 - \mu_1 - \mu_3 - \mu_5 - \mu_7$$

$$\lambda_{AB} = \mu_1 + \mu_2 + \mu_7 + \mu_8 - \mu_3 - \mu_4 - \mu_5 - \mu_6$$

$$\lambda_{AC} = \mu_1 + \mu_3 + \mu_6 + \mu_8 - \mu_2 - \mu_4 - \mu_5 - \mu_7$$

$$\lambda_{BC} = \mu_1 + \mu_4 + \mu_5 + \mu_8 - \mu_2 - \mu_3 - \mu_6 - \mu_7$$

$$\lambda_{ABC} = \mu_2 + \mu_3 + \mu_5 + \mu_8 - \mu_1 - \mu_4 - \mu_6 - \mu_7$$

The corresponding sample values of these contrasts may be represented by simply replacing μ by $\bar{X}$. For example,

$$L_A = \bar{X}_5 + \bar{X}_6 + \bar{X}_7 + \bar{X}_8 - \bar{X}_1 - \bar{X}_2 - \bar{X}_3 - \bar{X}_4.$$

A straightforward method for generating the coefficients necessary for these comparisons is given in Table 1.

Insert Table 1 about here

To generate these coefficients first write down those for the main effects. Place a plus one beneath each treatment where the main effect considered is at its upper level and a minus one beneath each treatment where that effect is at its lower level. The coefficients for the two way interactions are now obtained by forming the products of the coefficients for the main effects in each interaction beneath each treatment. For example, for each treatment A X B = AB. The three way interaction is dealt with in a similar fashion except that the product may be formed in any way as long as each main effect is used once and only once. For example, A X B X C = ABC, or AB X C = ABC, or C X AB = ABC, etc.

PRIOR DISTRIBUTIONS

As previously mentioned, when this solution is employed no restrictions are placed upon the prior distributions selected. An experimenter may use any that are dictated by the logic of the problem at hand. Of course, if a diffuse or a normal prior is assigned to a contrast, then the form of the posterior is immediately known for that contrast as pointed out by Lindley (1965). If, on the other hand, a prior of unusual form is employed, the posterior is still readily obtainable by rather straightforward numerical methods as shown by Powers (1973).

Generally the conceptualization of priors for main effects should not be troublesome for an experimenter. Those for interactions may be somewhat more difficult, but some aid with this chore can be derived by investigating the orthogonal contrasts. For example, Table 1 shows that, in terms of treatment combinations, as estimate of the AB interaction is given by

$$a_1b_1c_1 + a_1b_1c_2 - a_1b_2c_1 - a_1b_2c_2 - a_2b_1c_1 - a_2b_1c_2 + a_2b_2c_1 + a_2b_2c_2.$$

Algebraic simplification of this expression yields

$$(c_1 + c_2)\ (a_1 - a_2)\ (b_1 - b_2).$$

Thus a central point in the interaction can be located from estimates of the main effects. If a normal prior, for example, was to be assigned to this interaction, this point would be a logical choice for the mean of that normal distribution.

This procedure might be especially helpful in dealing with higher order interactions which are typically difficult to conceptualize. Here, for the ABC interaction one obtains

$$- a_1b_1c_1 + a_1b_1c_2 + a_1b_2c_1 - a_1b_2c_2 + a_2b_1c_1 - a_2b_1c_2 - a_2b_2c_1 + a_2b_2c_2$$

and

$$(a_1 - a_2)\ (b_1 - b_2)\ (c_2 - c_1).$$

LIKELIHOOD DISTRIBUTIONS

In a 2^k factorial arrangement we have 2^k treatment combinations. Assume that K observations are made in each treatment.

Thus $\bar{X}_i = N(\mu_i, \sigma^2/K)$ (this notation is read as $\bar{X}_i$ is distributed normally with mean μ_i and variance σ^2/K). Denote a linear contrast of these means as

$$L_s = a_{1s}\bar{X}_1 + a_{2s}\bar{X}_2 + \text{----} + a_{2^k s}\bar{X}_{2^k}.$$

Application of the linear function theorem to L_s now gives

$$E(L_s) = a_{1s}\mu_1 + a_{2s}\mu_2 + \text{----} + a_{2^k s}\mu_{2^k}$$

$$= \sum_{i=1}^{2^k} a_{is}\mu_i$$

and

$$V(L_s) = a_{1s}^2\left(\frac{\sigma^2}{K}\right) + a_{2s}^2\left(\frac{\sigma^2}{K}\right) + \text{---} + a_{2^k s}^2\left(\frac{\sigma^2}{K}\right)$$

$$= \sum_{i=1}^{2^k} a_{is}^2\left(\frac{\sigma^2}{K}\right)$$

$$= \frac{\sigma^2}{K}\sum_{i=1}^{2^k} a_{is}^2$$

$$= \frac{2^k\sigma^2}{K}$$

Thus when $\bar{X}_i \overset{d}{=} N(\mu_i, \frac{\sigma^2}{K})$, $L_s \overset{d}{=} N(\sum_{i=1}^{2^k} a_{is}\mu_i, \frac{2^k\sigma^2}{J})$.

POSTERIOR DISTRIBUTIONS FOR THE ORTHOGONAL CONTRASTS

The posterior distribution for each contrast in the orthogonal set is now derived by the direct application of Bayes formula;

$$P(\lambda_{js}|L_s) = \frac{P(\lambda_{js}\ L_s)\ L\ (L_s|\lambda_{js})}{\sum_{j=1}^{M} P(\lambda_{js})\ L\ (L_s|\lambda_{js})}$$

where $P(\lambda_{js})$ is the prior probability of the j th specified value of the s th orthogonal contrast.

$L(L_s|\lambda_{js})$ is the likelihood of the observed value of L_s, given the j th specified value of that contrast.

$P(\lambda_{js}|L_s)$ is the posterior probability of the j th specified value of the s th orthogonal contrast.

and M is the number of values for prior prior probabilities have been specified for that contrast.

POSTERIOR DISTRIBUTIONS FOR OTHER CONTRASTS

As pointed out by Powers (1971), if a set of orthogonal linear contrasts among the treatment means and a linear function of these means (mutually orthogonal to the contrasts) representing their sum is specified, then new linear functions of the original functions, which represent the individual contrasts among the means can be derived. Thus when prior distributions are specified on the original set of functions, priors have also been implicitly placed on the individual treatment means and on all possible Scheffé type contrasts of these means. Therefore, once the posterior distributions have been obtained for the original set, elementary rules of probability allow the derivation of the posterior for the new set.

Notice, in the discussion, that the terms "contrast" and "function" are used. This is because the functions specified are not always contrasts since $\sum_{i=1}^{2^k} a_i$ is not necessarily equal to zero. For example, in a 2^k factorial the function representing the sum of the treatment means is

$\lambda_T = \sum_{i=1}^{2^k} a_{iT}\mu_i$ where $a_{it} = 1$ for all i. Thus $\sum_{i=1}^{2^k} a_{iT} = 2^k$. A function representing a single treatment mean is $\lambda_m = \mu_m$ and $\sum_{i=1}^{2^k} a_{im} = 1$.

The first step in deriving the posteriors is to obtain the functions of the original set that represent the desired comparisons by evaluating the matrix equation

$$BC^{-1}\Lambda = \Lambda'$$

where

C is a k x k matrix of the coefficients of the original orthogonal set and each row of C contains the coefficients of one function.

B is an m x k matrix of the coefficients for m new functions representing individual treatment means and other contrasts, and again each row of B represents the coefficients of one function. (These functions need not be orthogonal to themselves nor to the orthogonal set.)

Λ is a k x k column vector of the orthogonal functions.

and Λ' is a column vector of the new functions.

The posterior distributions for the functions in Λ' are now obtained by forming the proper combinations of the probabilities in the posterior distributions for the functions in Λ.

CONCLUSION

The method presented in this paper is a relatively straightforward and practical procedure for conducting a Bayesian analysis of a 2^k factorial arrangement of treatments. If an experimenter employs one of these designs, he probably is most interested in, and has most prior information about, the main effects and interactions. This solution allows prior distributions to be placed on, and posteriors derived for, these most salient effects. It also permits further analysis of interactions, where desired, through the use of Scheffé type contrasts. In addition, the calculations required are not unduly cumbersome regardless of the forms of prior distributions chosen.

Lastly, this solution can be applied to factorials other than the 2^k series with only slight modification.

1. Coefficients for a Set of Orthogonal Contrasts of Treatment Means Representing Main Effects and Interactions in a 2^3 Factorial Arrangement of Treatments

in Effect or teraction	Treatment Combination							
	$a_1b_1c_1$	$a_1b_1c_2$	$a_1b_2c_1$	$a_1b_2c_2$	$a_2b_1c_1$	$a_2b_1c_2$	$a_2b_2c_1$	$a_2b_2c_2$
	Mean of Treatment Combination							
	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7	μ_8
A	-1	-1	-1	-1	+1	+1	+1	+1
B	-1	-1	+1	+1	-1	-1	+1	+1
C	-1	+1	-1	+1	-1	+1	-1	+1
AB	+1	+1	-1	-1	-1	-1	+1	+1
AC	+1	-1	+1	-1	-1	+1	-1	+1
BC	+1	-1	-1	+1	+1	-1	-1	+1
ABC	-1	+1	+1	-1	+1	-1	-1	+1

REFERENCES

Box, George E. P., and Tiao, George C. (1968), "Bayesian Estimation of Means for the Random Effect Model," Journal of the American Statistical Association, 63, 174-181.

Duncan, David B. (1961), "Bayes Rules for a Common Multiple Comparisons Problem and Related Student - t Problems," Annals of Mathematical Statistics, 32, 1013-1033.

_____ (1965), "A Bayesian Approach to Multiple Comparisons," Technometrics, 7, 171-222.

Hughes, Francis P., and Powers, J. E. (1971), "On Replacing the Population Variance by its Sample Estimate in Bayesian Analysis of Planned Orthogonal Comparisons," Experimental Publication System of the American Psychological Association, 10, Ms. # 352-24.

Lindley, Dennis V. (1965), Introduction to Probability and Statistics from a Bayesian Viewpoint, Cambridge, England: Cambridge University Press

_____ (1972), "A Bayesian Solution for Two-Way Analysis of Variance," American College Testing Program, Technical Bulletin #8, Iowa City, Iowa.

_____, and Smith, A.F.M. (1972), "Bayes Estimates for the Linear Model," Journal of the Royal Statistical Society, Series B, 34, 1-41.

Powers, James E. (1971), "A Posteriori Comparisons: A Bayesian Procedure," paper presented at the Northeastern Educational Research Association in Ellenville, New York.

_____ (1973), "A Bayesian Analysis of Planned Orthogonal Comparisons," Journal of Experimental Education, 42 (2), 73-79.